

University of Groningen

The Good, the Bad, and the Brain

Schirmann, Felix

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2014

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Schirmann, F. (2014). *The Good, the Bad, and the Brain: Theory and History of the Neuroscience of Morality*. [Thesis fully internal (DIV), University of Groningen]. [s.n.].

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 3

The neuropathology of morality: Germany 1930–1960

Published as:

Schirmann, F. (2014b). The neuropathology of morality: Germany 1930–1960. *Journal of the History of the Neurosciences*, 23(1), 56-74. doi:10.1080/0964704X.2013.804637

Abstract

This article analyzes brain scientists' attempts to trace morality in the brain in Germany from 1930 to 1960. The debate around Karl Kleist's localization of the *Gemeinschafts-Ich* [community-I] in the 1930s is depicted in order to illustrate the central arguments for and against localizations of morality. The focus of this article is on the period 1936–1960 in which experts put forth specific ideas on morality's cerebral underpinnings that mirror the larger theoretical shift from strict localization doctrine to a more holistic understanding of the brain. As a result of this shift, experts avoided exact localizations of morality. Instead, they posited correlations between brain areas and morality. The analysis illustrates the dependence of neuropathological research on morality on general theories of brain functioning and marks a first contribution to the history of the neuroscience of morality for the time after 1930.

Introduction

Modern neuroscience ventures to trace morality in the brain since around the year 2000. Neuroimaging technologies offer glimpses into the brains of moral agents and neuroscientific evidence on criminals' brains features in court cases (Fumagalli & Priori, 2012; Hughes, 2010; Sinnott-Armstrong, 2008). However, brain science's quest to study morality is much older. Throughout the last 200 years, brain researchers made various attempts to describe, map, and dissect morality, yielding speculations on moral centers, criminals' aberrant cytoarchitecture, and postencephalitic moral insanity (Verplaetse, 2009). This line of research was subject to controversy in the scientific community. To many experts, morality, along with other higher mental functions, lay beyond the capabilities of brain science. Consequently, there were disputes over research methods, the validity of data and the proposed neuronal basis of morality. For example, at the end of the nineteenth century, experts assailed and even ridiculed colleagues who proposed localizations of morality (e.g. Verplaetse, 2004). Thus, the history of the brain science of morality provides an intriguing example of brain science's borderland and scientists' boldness.

In the past, the brain science of morality was marginal; just as its history. To the best of my knowledge, aside from Verplaetse's (2009) seminal work which starts its coverage around 1800 and fades out in the 1930s, historians of science have not explicitly dealt with the topic. Verplaetse identified a "rupture" (p. 27) – a major transition in scientific thinking – around 1930. This "crisis" (p. 27) was characterized by increased skepticism regarding neurobiological views on morality and an associated take-over of psychologists and sociologists, who framed morality as a social phenomenon. As a consequence of this transition, brain science of morality became even rarer. However, the period after 1930 remains unaddressed. In order to offer a first contribution to this neglected topic in the history of neuroscience, this paper analyzes neuropathologists' attempts to study morality in the brain in Germany in the

years 1930 to 1960. In this period, several experts searched for cerebral correlates of moral behavior, studied moral lapses after frontal brain lesion or posited the existence of brain-based disorders of decency and ethical ability. This research coincided with a more general debate on brain functioning. Old-fashioned localization doctrine – that is, the idea that the brain is organized in distinct modules corresponding to specific functions – was challenged by holistic understandings that stressed the integrated functioning of the whole brain (Finger, 1994; Harrington, 1996). Correspondingly, attempts to trace morality in the brain at that time reflected the tension between these antagonistic theories (see below).

In this paper, I show that there were efforts to study morality in the brain after 1930. In order to set the stage for this, I describe the debate following Karl Kleist's localization of the *Gemeinschafts-Ich* [community-I] in the 1930s, scrutinize the criticisms, and analyze the ensuing dispute that echoed through the following decades. The focus of this paper is on the period 1936-1960 in which experts put forth specific ideas on morality's cerebral underpinnings that mirror the larger theoretical shift from strict localization doctrine to a more holistic understanding of the brain. As a result of this shift, experts avoided exact localizations of morality. Instead they posited correlations between brain areas and morality. I analyze the neuropathology of the moral brain in light of this transition by depicting the surrounding debate. This analysis illustrates the dependence of attempts to trace morality on general theories of brain functioning and offers a first contribution to the history of the neuroscience of morality for the time after 1930.

The remainder of this paper is structured as follows: In order to constitute the historical background, the situation in German brain science and brain-based morality research around 1930 is depicted. Special attention is paid to the localization doctrine as method and theory. Then Kleist's localization of morality is presented and the ensuing debate analyzed in order to highlight the major objections against such

attempts. As a next step, researchers' efforts to study morality in the brain in the period from 1936 to 1960 are described and contextualized with the more general debate on brain functioning. In conclusion, the transformation from localizing moral centers to conceiving of brain areas as being related to morality is analyzed for the respective period.

Politics, German Brain Science, Localization Doctrine, Morality: The Situation around 1930

The political and scientific climate in Germany changed dramatically several times between 1930 and 1960 (Vom Bruch & Kaderas, 2002). In these times of upheaval, German brain scientists had to deal with Nazi politics and race ideology, the expulsion of Jewish and dissident colleagues, inadequate research conditions during the Second World War, and later isolation from their international colleagues (Peiffer, 1998; Peiffer, 2004). From 1933 to 1945, large-scale ties between the Nazi regime and neurology and psychiatry existed (Schmuhl, 1987; Weindling, 1989). The Nazi Party's political and private control restricted scientific freedom and demanded allegiance of scientists. Despite these constraints, brain science continued throughout the war and there was room to maneuver for scientists (Roelcke, 2008; Roelcke, 2010).

Along with the political developments came a rise and subsequent decline of biological theories for behavior, mental disorder, and alleged inferiority. Neuro-/biological theories had been in demise since around 1900 (Shorter, 1997), but during the period of Nazi rule, they experienced an unprecedented popularity. Claims about inferior biology were central to Nazi race ideology, eugenics, and bio-criminology (Schott, 2002; Wetzell, 2000).⁷ With the end of Nazi rule, the race-based views vanished (Roelcke, 2005). Aside from these external developments, an internal battle between neurologists and psychiatrists had divided the scientific community in the

years after the turn of the centuries (Engstrom, 2003; Roelcke, 2002). The significance of the brain for mind and behavior was among the topics that caused continuing disagreement. Psychiatrists were skeptical of brain-based explanations for mental disorder and sought independence from the neurologists. In 1935, the Nazi rulers ordered the establishment of a joined society for both, but the tensions between disciples of the two disciplines did not abate.

In the second decade of the twentieth century, the First World War brought numerous patients with head wounds under the eyes of brain scientists. The abundance of cases boosted neuropathology. The brain injuries were manifold: degree, location, and associated symptoms varied considerably. As a consequence, a plethora of patients with diverse lesions and psychological dysfunctions, along with a multitude of expert's opinions formed a highly-confusing puzzle. Yet, seminal neuropathological work was based on these cases in the interwar years (Feuchtwanger, 1923; Poppelreuter, 1917; von Schjerning, 1922/1934). This work was largely predicated on the localization doctrine.

The Localization Doctrine: Characteristics, Criticisms, and Changes

Localization doctrine played a dominant role in brain scientists' theorizing since the time of the great localizations (e.g. speech center and motor cortex) in the second half of the nineteenth century. However, the doctrine was contested since its inception. Especially the localization of higher mental functions spurred scientists' disagreement (Finger, 1994; Hagner, 2008; R. M. Young, 1970). To understand the debate depicted in the remainder of the paper a closer look at the localization doctrine – as a theory and a method – is necessary.

The key tenet of localization doctrine as a theory consisted in the “fragmentation of the psyche” (Hagner, 2008 p. 20) or “decomposition” (Riese & Hoff, 1950 p. 63) – that is, the division of the psyche into independent units. It was assumed that these psychological units could be assigned to circumscribed neuroanatomical counterparts, with a differing degree of rigor ranging from correlation to strict correspondence. Accordingly, several versions of the localization doctrine co-existed that varied regarding their spatial rigidity. The strictest and most controversial version posited “‘punctuate’ or ‘point-to-point’ localization of nervous function” (Riese & Hoff, 1951 p. 452) – that is, the exact matching of certain psychological capacities with distinct neuroanatomical centers. As a method, the localization doctrine proceeded by observing psychological dysfunction after brain damage or brain stimulation and charting the respective cerebral location. Several approaches helped to pinpoint brain regions: Staining methods to identify particular tissue in autopsy, electro-stimulation of certain regions, induced brain ablations in animals, and the observation of specific psychological dysfunction after lesions in humans. Generally, the observation of a coincidence of a defined lesion with a specific dysfunction was central. Based on the location of the lesion, the cerebral location of the (unimpaired) psychological function was then inferred. As Tizard (1959 p. 143) put it: “A lesion at X produces symptom T. This represents damage to the psychological trait Z which must be localized at X.” Interwoven in this “logic of deletion” (Star, 1989 p. 12) is the primacy of spatial information over, say, temporal information (Riese, 1967). Moreover, the design of localization theory rested on this reasoning method and its implicit assumptions: Circumscribed brain areas, modularity of psychological capacities, associated distinct malfunction, and the combining inference structure formed a stencil upon which localization theory was modeled. Close ties existed between the tool and the theory (see Gigerenzer, 1991). In summary, two claims were inherent in the localization doctrine: an ontological claim

concerning the existence of distinct brain centers representing specific psychological functions and an epistemic claim regarding the reasoning from lesions to intact centers.

A specific objection against the localization doctrine surfaced at the beginning of the twentieth century in German-speaking brain science. Constantin von Monakow and Kurt Goldstein assailed the crudeness of the doctrine and called for a reformulation. According to them, the brain functioned as an integrated whole rather than as a bundle of isolated units. This holistic position emerged as an influential alternative theory of brain functioning in the German debates after 1900 (Isler, 2010; Harrington, 1996; Riese, 1967; see also Ash, 1995). Although the localization doctrine in general was challenged, extenuated versions of it persisted. For example, Hans Berger (1923) endorsed strict localizations, but along with others he questioned the rigidity of the divisions between the brains' putative centers. The associated debate on the validity of the doctrine ranged from the microscopic (cytoarchitecture) to the macroscopic level (large brain areas) (Peiffer, 2004). Despite the growing controversy, the doctrine was still conceived of as a prolific research strategy for the future (e.g. Spielmeier, 1930).

Traces of Morality in the Brain

While the debate on the localization doctrine smoldered around 1930, some brain scientists searched for traces of morality in the brain. Relating human capabilities to brain composition was not uncommon at the time. For example, a current in brain science studied the brains of the elite and criminals (Hagner, 2004). Contemporary researchers followed multiple tracks of brain-based morality; not only in Germany: British, American, and French experts made efforts to understand the neuro-/biology of good conduct (Richards, 1987), to identify putative moral centers (e.g. Browning,

1921; Yawger, 1935) or to research the conspicuous ethical breaches evinced by patients with Encephalitis lethargica (Verplaetse, 2009). The various experts used the term morality (or cognates) ambiguously. It could refer to outright criminal offences, exuberance, disinhibition, tactlessness, egotism, a lack of remorse, or minor indecency. The diversity of socially deviant behaviors studied as neuropathological or bio-medical phenomena was striking.

In Germany, several scientific approaches to link deviancy, immorality, or criminality to biology in general or the brain in particular co-existed. Bio-criminologists studied the heritability of criminal tendencies, leading to the controversially discussed sterilization and killing of criminals during the Nazi regime (Wetzell, 2000). Psychiatrists researched the genetics of mental disorders accompanying immorality, such as moral feeble-mindedness [moralischer Schwachsinn] (Dubitscher, 1935), Kurt Schneider's compassionless [gemütlos] psychopathy, or Karl Birnbaum's amoral psychopathy (Werlinger, 1978). Certain experts speculated on a cerebral basis for moral insanity; a disorder marked by moral lapses without impairment of the intellect.⁸ Yet, only few brain injured showed “severe character changes in terms of ‘moral insanity’ ” (Scheid, 1939 p. 270). Although this controversial disorder was becoming outdated by the 1930s, the debate on it highlighted a significant distinction: Immorality as congenital mental disorder was different from acquired immorality resulting from brain damage (Glaser, 1932). How did neuropathologists account for immorality after traumas, lesions, or tumors?

Experts knew since the nineteenth century that brain damage could lead to undesirable character changes. Tumors or trauma to the frontal lobes were associated with disinhibition, euphoria, indecency, and apathy (Finger, 1994). For example, Welt (1888) suggested a connection between indecent and immoral behaviors and destruction of the right orbital region. Several other researchers documented cases of patients' immorality after frontal brain damage (Feuchtwanger, 1923; Grünthal, 1936;

Pfeifer, 1928a; Pfeifer, 1928b; Ruffin, 1939; Schuster, 1902). However, hardly any researcher pinpointed morality's cerebral position. The case material was too heterogeneous, the locations of the suspicious brain damage varied, and only some patients evinced immoral conduct. Despite this lack of clarity, one prominent brain scientist suggested a localization of morality.

Kleist's Localization of Morality and its Critics

Karl Kleist was Professor of Psychiatry in Frankfurt and an important figure in German brain science. His research on neuropathology and his neuropsychiatric system were seminal (Neumärker & Bartsch, 2003). Kleist (1934) had written a standard work on brain pathology⁹ based on macroscopical considerations of brain injuries from the First World War. In it, he localized *Funktionskreise des Ichs* [functional circuits of the I]. He identified a *Gefühls-Ich* [Emotional-I], a *Trieb-Ich* [Drive-I], a *Körper-Ich* [Body-I], a *Selbst-Ich* [Self-I], a *Welt oder religiöses-Ich* [World or Religious-I], and a *Gemeinschafts-Ich* [Community-I]. He added that “each can sicken on its own, and thus must represent a physiological unit, which has distinct anatomical carriers” (p. 1169). Most important for the analysis here, the community-I corresponded to “fidelity, gratefulness, helpfulness, friendship and hostility, mistrust and spitefulness, need for recognition, imperiousness and obedience” and the “ethical attitudes and their expression” (p. 1169). Kleist localized the community-I in both sides of the orbital brain¹⁰ and thus assigned a cerebral position to morality.

Kleist presented his work on the structured I at a meeting of the German Psychiatric and Neurological Society in 1936 (Kleist, 1937a).¹¹ The reception of his talk was caustic. His colleagues assailed the notion of localization in general and Kleist's reasoning method in particular. Friedrich Quensel, Professor for Psychiatry and Neurology in Leipzig, commented: “To find a localization [...] for higher mental

functions is much harder, almost impossible, especially if one considers how incomplete and vague our knowledge and notions of the nature and the limits of higher mental functions generally is” (Miscellaneous, 1937 p. 331). Many attendees shared this distrust in the capabilities of neuropathology and considered Kleist's attempts as premature. Moreover, the commentators voiced scathing criticism against Kleist's reasoning. Johannes Lange, Professor of Psychiatry in Wroclaw, cited Hughlings Jackson's famous distinction between a localization of symptoms and a localization of functions (“To locate the damage which destroys speech and to locate the speech are two different things.” (cited in Lange, 1937 p. 250)); Lange accused Kleist of conflating symptoms with functions. It was bad logic to reason from a localizable psychological dysfunction (a negative) to the location of a psychological function (a positive). Accordingly, one cannot infer the existence of centers. This ‘negative argument’ reverberated through many other comments. The “logic of deletion” (see above) was flawed. Oswald Bumke, Emil Kraepelin's successor on the Chair for Psychiatry in Munich, claimed that if one proceeded according to Kleist's argumentation one could reason from the cutting of the optic nerve to the location of the visual center in the optic nerve itself; which was evidently preposterous. In addition, the arbitrary rigidity of Kleist's categorization of the I did not fare well with his colleagues. In his second comment, Bumke objected to Kleist's reification of the structured I: “I cannot get rid of the impression that one wants to play Tennis with soap bubbles” (Miscellaneous, 1937 p. 330). Hugo Spatz (1937), on the verge of becoming Director of the Kaiser-Wilhelm-Institute for brain research in Berlin, was concerned regarding Kleist's “dissection of the psyche” (p. 230). Spatz endorsed the idea of localizations of psychological functions. He had witnessed striking character changes in a patient with destroyed orbital lobes (Bostroem & Spatz, 1929) and in others with Pick's disease. According to Spatz, particularly the destruction of the frontal pole of the neocortex (Brodman area 10, 11, and 47) disturbed the “highest

mental functions” (Spatz, 1937 p. 212). In addition, the described region was the phylogenetically youngest part of the brain and thus a candidate for the specifically human functions. Although Kleist received some support from Spatz, a majority of the attendees rejected Kleist’s ideas.

Kleist retorted: “How can one consider me capable of such nonsense?” (Kleist, 1937b p. 340). He disavowed that he was a localizer and emphasized the tentativeness of his findings. In this context, Kleist identified with an extenuated version of the localization doctrine. He suggested a “connection of certain functions and malfunctions not to specific places, but [to] partly large-scale or widespread formations of the nervous system” (p. 337) and adhered to a “general biological principle of division of labor in the nervous system, [that is,] the allocation of diverse functions to diverse organ parts” (p. 337). So there were no centers and no localization of morality? It can be doubted that Kleist meant what he said. The wording in his publications suggests that he was a staunch believer in strict localization doctrine (see Kleist, 1934; Kleist, 1935; Kleist, 1937a). Furthermore, Walther Riese, Historian of Neurology and former student of Kleist, considered Kleist to be one of the last true localizers (Riese, 1959; see also Uttal, 2001).

After the disappointing conference, Kleist continued his impressive career (Neumärker & Bartsch, 2003). His ideas on the structured I featured popularly in the following decades’ publications. Although his localization of morality attracted little attention in the years to come, it still inspired reverie and fruitful follow-up work as well as more criticism. In a letter, Kurt Schneider, then Head of the Clinical Department at the Kaiser-Wilhelm-Institute in Munich, wrote to Karl Jaspers, Philosopher and Author of a standard work on psychopathology, it is stated that “[o]ne is allowed to say that Kleist’s localization theory stands aside from clinical psychiatry. [...] It is a true brain mythology” (see Peiffer, 2004 p. 569).¹² Were strict

localizations of higher mental functions decried as mythology in the decades to come? And was morality in the brain just another myth?

Tracing morality in the brain: 1936-1960

In 1936, Hertha Cosack, who worked with Lange in Breslau, published a paper titled “criminal personality change due to frontal brain damage” (Cosack, 1936). In a suicide attempt, her patient had shot himself through the forebrain from right to left. His relatives reported alarming personality changes, consisting in moral lapses, indecency, and callousness, yet no intellectual deficits. A year later, Lemke (1937) explained the brazenness of one of his patients with a tumor on both sides of the orbital brain as a “dysfunction of the community-I” (p. 65). Though rare, these character changes after impairment of the orbital brain were acknowledged as a specific neuropathological syndrome (see also Pittrich, 1941). The patients “become childish, superficial, silly, tend to scoffing and joking, are erratic and weak [haltlos], often also egoistic, spiteful, indecent and shameless, take liberties with morality, are agitated and occasionally criminal” (Duus, 1939 p. 646). Another expert summarized the patients’ transformation as follows:

“Sick persons with injuries and tumors in the orbital brain provide the most glaring examples of acquired callousness [Gemütlosigkeit] loss of all finer moral, social and esthetic impulses, of raw unrestrained drive [Triebhaftigkeit]. Previously inconspicuous people conform to callous [gemütlos] psychopaths, cases of 'moral insanity', lose shame and sense of honor, compassion, consideration for others, turn indecent, dishonest, neglect their duties, commit crimes.” (von Baeyer, 1941 p. 25)

These shifts in moral behavior had been known for several decades (see above). Then, Welt (1888), Kleist (1934) and others had categorized these phenomena by using versions of the localization doctrine. Did researchers in the 1930s follow suit?

Maximinian de Crinis, convinced National Socialist and Head of the Psychiatric Clinic in Berlin (see H. Jasper, 1991), proclaimed that “that, what we name with the often incorrectly used word morality, is also a function of the brain” (de Crinis, 1938 p. 426). Yet, he was skeptical of Kleist’s idea that morality could be pinpointed in the frontal brain. First, he challenged that lesions to the lower part of the frontal brain yielded immorality in all cases. Second, he claimed that the brain areas responsible for “*Gesittung*” [civilized mode of behavior; morality] (p. 435) were dispersed over the entire brain. Thus, de Crinis rejected the determinism associated with the localization doctrine as well as its notion of discrete brain centers. Joachim-Ernst Meyer (1940), a staff member of de Crinis, agreed. Although he acknowledged the existence of “orbitofrontal deficit beings”¹³ (p. 371) and their ethical breaches, he believed that an exact localization of morality was impossible. However, an identification of the different neuronal elements on which morality depended, such as the interbrain, could be achieved. In this view, morality originated in several brain regions. The frontal brain as Kleist’s main candidate for morality began to lose brain scientists’ confidence.

At the time, the frontal brain was an object of dispute (Filley, 2010). Brain scientists still analyzed the cases from the First World War (Scheid, 1939) and tried to account for the “myriad of deficits” (Finger, 1994 p. 325) that was associated with lesions to this region. A majority of the experts agreed that destruction of the frontal brain rarely led to deficits in intelligence and memory. However, experts observed diverse transformations in impulse, mood, temperament, and personality (Rylander, 1939). Incomprehension abounded. In a review on the frontal brain syndrome, Ruffin (1939) listed the diverse psychological dysfunctions ranging from euphoria, moria or *Witzelsucht* [joking addiction], hypomania, depression, apathy, to ethical inability.

Corresponding to this wealth of psychological impairments was the elusive statement that the frontal brain had “a particularly human function” (Ruffin, 1939 p. 58). A proposition that accounted for the complexity came from Goldstein (1934), who suggested an impairment of the general ability to abstract in frontal lesion patients. Others declared the frontal lobes to be the source for general drive (e.g. Beringer, 1941). These diverging propositions show how far from being settled the matter was: Cases were abundant and the psychological dysfunctions diverse. The number of brain injured further increased with the Second World War. After it, 150,000 brain injured were registered in Germany (Lindenberg, 1950). Again, this case material proofed invaluable for brain scientists and spurred new interest in brain pathology and localization questions (Bay, 1947; Heygster, 1949).

In addition to the war cases, the practice of leucotomy produced numerous patients with instructive lesions. Though never a widespread therapeutic tool in Germany, German brain scientists frequently commented on the practice, its neuroanatomical fundaments, and the lessons learned from it. The surgical intervention and the brain areas destroyed varied in leucotomy, but the majority involved ablating parts of the frontal brain or its connections (Hassler, 1950; Valenstein, 1986). Some surgeons specifically targeted the orbital brain (e.g. Knight & Tredgold, 1955; Ström-Olsen & Northfield, 1955). If Kleist was right, then destroying this area was tantamount to destroying morality. Though Kleist himself objected to psychosurgery (cf. Valenstein, 1986 p. 217-218), other researchers used his localization of the structured I to make sense of the psychological effects of leucotomy in anatomical terms (Barahona Fernandes, 1950; Walch, 1956). Furthermore, surgeons reported alleviation of psychotic symptoms after operating on the orbital brain, yet side-effects, among them apathy, lack of self-criticism and loss of ethical inhibition occurred occasionally. The character changes in the leucotomized resembled the moral lapses of the orbital lesion patients and were used as an

argument against this anatomical therapy (Haddenbrock, 1949; Haddenbrock, 1952; Stoll, 1954). Operating on the frontal brain to some was “a crude almost childish endeavor” (Glees, 1949 p. 220).

In general, character changes after lesion formed the main line of evidence in attempts to trace morality in the brain around the middle of the twentieth century in Germany – with one exception. In 1951, Oskar Vogt, then aging doyen of brain science and beyond his zenith, published a paper addressing cytoarchitectural differences in murderers. According to him, immorality expressed itself in a peculiar cell alignment in the lamina pyramidalis of the cortex.¹⁴ However, this type of microscopical analysis was not used widely in attempts to uncover morality’s cerebral underpinnings. Rather the neuropathological method, as Kleist had applied it, formed the basis for further publications on morality in the brain in the 1950s. Objections to this method and its theoretical basis were frequent.

Researchers had registered inconsistencies and a lack of exactitude in localization questions since the 1930s. The goal of a precise mapping of higher mental functions on distinct anatomical counterparts repeatedly failed and researchers more and more curbed their aspirations. For example, Eberhard Bay, later to become Professor of Neurology in Düsseldorf, commented “pushing precision in brain pathology too far necessarily leads to pseudo-exactitude” (1947 p. 300). Bay repeatedly attacked the localization doctrine as a methodological tool and as a theory of brain functioning (Bay, 1941; Bay, 1947; Bay, 1951; Bay, 1953). To him, the existence of areas for higher mental functions was an axiom, not the result of empirical tests. Just because motor functions could neatly be mapped on the motor cortex, the inference by analogy that such circumscribed areas existed for other psychological functions did not automatically hold true. Thus, deficits after lesions had to be interpreted differently: A lesion in the frontal areas did not destroy an isolated center but disturbed the interplay of the entire brain. Applied to morality this

meant that immoral patients did not lack a specific moral brain area and that morality was scattered over the whole brain. Acknowledgment of these objections was visible in research on brain-based morality in the 1950s (see below).

Ernst Kretschmer, Professor in Tübingen (see Priwitzer, 2007), addressed morality in the brain on multiple occasions. In the 1920s, Kretschmer (1921) had devised a psychopathological system based on different body types. For each body type, the interplay of psychological factors and biological constitution yielded specific characteristics and disorders. In later versions of the respective monograph, Kretschmer posited relations between bodily constitution and crime (Kretschmer, 1955). A sideline in his argument rested upon observations on patients with lesions in the orbital and interbrain. In 1949, Kretschmer dedicated a paper to Kleist in which he formulated the concepts “*sphärische Integrierung*” [spherical integration] and “*dynamische Steuerung*” [dynamic control] (p. 470). It described the integration of pre-conscious, inhibitory, and general-affective representations of a given situation with the action required in that situation. This integration was disrupted in orbital brain patients, leading gradually to indecency and in severe cases to “loss of ethical control” (Kretschmer, 1949 p. 471). Moral behavior here had a twofold source: the orbital brain and the interbrain. Kleist himself had speculated on the significance of the interbrain for character (Kleist, 1934). In a later publication titled “The social and moral defect as biological problem”, Kretschmer maintained that specific kinds of ethical disinhibition as a result of unchecked drives originated only from isolated lesions in certain brain locations – that is, orbital brain, interbrain, and brain stem (Kretschmer, 1953a). He expounded “[t]he majority of these drive disorders lead to contact disorders with the environment, sociological problems but also to criminal lapses” (p. 32). In two other publications, Kretschmer elaborated on the link between brain lesion and criminality: Minute brain damage evaded the eye of the untrained physician and could precipitate criminal behavior, resulting in a misinterpretation of

the brain injured as merely criminal. Thus, neuro-psychiatric knowledge was essential for the legal assessment of criminals (Kretschmer, 1953b; Kretschmer, 1954). Despite the proclamation of clear-cut connections between brain damage and immoral behavior, Kretschmer was not an old-fashioned localizer. Localization to Kretschmer signified a division of labor in the brain, but not the existence of distinct centers. Furthermore, he understood localization as a statistical problem with correlations between lesions and dysfunction showing only on a large-scale data level (Kretschmer, 1956). This statistical understanding of localization was predicated on the observation that individual brains differed. Anatomical variation was not an anomaly and direct comparability of brains not a fulfilled precondition. Thus a prerequisite for generalizing with the “logic of deletion” was not given. A contemporary of Kretschmer, Schaltenbrand (1950), emphasized brains’ dissimilitude and advocated a statistical approach to localization. In his brain maps, he identified the frontal area and the tip of the temporal cortex as parts of a bipartite “*Ich-Hirn*” [I-brain]. Though Schaltenbrand did not pinpoint morality in the brain, he maintained that lesions to the frontal I-brain caused aimlessness, emotional blunting, a lack of self-criticism and suggestibility.

The statistical approach to localization was a reaction to a persistent problem: Variability. Not all orbital lesion patients changed in character and only a few turned immoral. Experts’ descriptions and diagnoses differed. Hans Walter Gruhle, Director of the Psychiatric Clinic in Bonn and an outspoken opponent of the localization human social functions and the ideas of Kleist and Kretschmer, commented on the verbal confusion:

“If one reads: failure of the community-I emanates from the interbrain, failure of community ethos emanates from the orbital cortex – or: only ethos is of cortical nature, while friendship and hostility have their basis in the interbrain,

there the author conjures up completely unexplained notions, which stem from popular psychology and which can only breed misunderstanding.” (Gruhle, 1948 p. 218)

Gruhle’s attack against imprecise description of symptoms appeared justified (see also Gruhle, 1954). The descriptions and the symptoms themselves were indeed diverse. Yet consistent enough for some researchers to propose an independent syndrome – consisting in disinhibition, ruthlessness, and aggression after orbital brain lesion – dubbed “posttraumatic organic character change” (Faust, 1955a p. 1238). Why this certitude amongst the ambiguity? Statistically orbital lesions patients committed crimes more often than other lesion patients (Hoheisel, 1954; Lindenberg, 1950; Rehwald, 1956). In an analysis of 6600 lesion cases, Geller (1960) found that 14 per cent of the orbital lesion patients came into contact with the law. This was 10 per cent higher than the next group of lesion patients. Geller underscored that there was no “clear parallelism” between “anatomical and psychological changes” (p. 404); hence, no exact localization. Yet, the statistical results proofed reliable enough for correlational ties between orbital brain lesion and criminality. This large-scale, statistical understanding of localization questions exemplified a new way to conceive of the ambiguity of neuropathological results. Inconsistent results had continuously posed a problem to the localization doctrine.¹⁵ For example, researchers tried to replicate Kleist’s localization of the *Gemeinschafts-Ich* with varying success, resulting in disbelief in the exactitude of localizations (Faust, 1955b; Heygster, 1949). Whereas experts in earlier times had quarreled over the implausibility of the results, brain scientists around the middle of the twentieth century presented this ambiguity as correlations: A lesion led to a dysfunction with a certain probability. This correlational understanding of the lesion-dysfunction-relation accommodated for the often found inconsistencies regarding morality’s neuronal basis.

Researchers more and more questioned the rigidity of Kleist's localization and adapted his theory in the light of novel knowledge about the brain. Although in the 1950s, his ideas still served as a guidepost. Hoheisel (1954) elaborated on Kleist's localization and Kretschmer's warning of undetected brain damage in criminals. He referred to a murderer in the 1930s who had suffered brain trauma during an accident. The murderer was executed and the autopsy revealed lesions in the right temporal, frontal and orbital brain. Hoheisel speculated that these lesions were responsible for the character changes that culminated in murder. The propensity of the brain injured to commit crimes was well known. According to Hans-Jochen Thelen, a staff member of Kretschmer in Tübingen, specific lesions corresponded with typical crimes: "So typical that we are almost tempted to speak of the act of a frontal-orbital- or interbrain injured" (Thelen, 1953 p. 237). Indeed, the crimes of these patients were special. Most often they were spontaneous, random, publicly conducted and readily admitted in interrogation. The simplicity and aimlessness of the crimes was striking. All in all, this indicated a lack of control or inhibition rather than callous, premeditated criminality (Lindenberg, 1950; Lindenberg, 1953). Still, legal responsibility became an issue: "The psychopathology of the frontal brain, particularly of the orbital brain, has always had an exceptional position in forensic assessment because lesions there lead to character changes that are especially suited to induce conflict with the law" (Leonhardt, 1956 p. 501). The "*Bund der Hirnverletzten*" [alliance for the brain injured] specifically endorsed the practice of consulting physicians' professional opinion in law suits against brain injured soldiers from the wars. According to Paragraph 51 of the German Penal Code¹⁶, offenders were not legally responsible if they did not have insight into the illicit nature of their action or if they were unable to act upon that insight due to impaired consciousness, pathological disorder of mental action, or mental weakness (Ewald, 1959; Kroeber, 2007). On these grounds, courts frequently exculpated brain injured (e.g. Thelen, 1953).

Although Kretschmer, Schaltenbrand, Hoheisel, Thelen, Lindenberg, and Leonhardt knew that destruction of the orbital brain frequently lead to immoral deeds, they did not infer a moral brain area from that. The strict inference from lesion to positive function was no longer justifiable as additional knowledge about brain functioning rendered this reasoning dubious. Generally speaking, researchers reworked the connection between orbital brain damage, character changes, and immorality in the 1950s, adding nuances in understanding the cases. Thelen (1953) discussed one of these nuances in his paper on “brain injury and criminality”. One of his patients suffered from a persistent “destruction of the orbitally represented community-feelings” (1953 p. 237). Yet, in several of his other patients, he had observed a recovery of the disturbed psychological functions after a certain amount of time. He attributed this to the brain’s capacity to compensate for damage. Other brain areas could take over the functions of those destroyed and in the course of time the psychological impairments diminished or vanished completely. The temporal dimension also played a decisive role in the assessment of the lesions. As time progressed, the symptoms of the brain injured changed. Symptoms evident shortly after brain trauma differed from those in later stages. Thus, the assessment of the lesion and the associated psychological dysfunction depended on the point in time – and not exclusively on the location. Temporal information appeared to be just as important as spatial information. Hence, a careful consideration of the course of the impairment was necessary (Faust, 1955b). Neuronal compensation and the significance of the temporal dimension in the assessment of lesions had been known since the beginning of the century (Riese, 1967), but it took several decades until this knowledge was properly integrated in the canon of brain science. The brain according to these qualifications is dynamic. This debilitated the classical “logic of deletion” since it reasoned from dysfunction to location without paying sufficient heed to time. There was no room for strict localizations of morality in a versatile brain.

Walch (1956) pointed to another peculiarity. He approved of Kleist's identification of the community-I and emphasized the significance of the interbrain for moral control. At the same time, he stressed that the neuronal sources of character were dispersed over the entire brain. Accordingly, Walch did not conceive of the orbital brain as a center for morality, but assigned a guiding, inhibitory function to it. Here, an abstract function replaced a moral center. Moreover, Walch emphasized that the appearance and severity of the orbital brain syndrome depended on the premorbid personality. Hence, a brain lesion only exacerbated immoral character tendencies that had existed before. Again, pointing to prior wickedness was not completely new. In the 1930s, Schneider (1935), Duus (1939), and Scheid (1939) had called attention to the interplay between lesion and personality. These calls were amplified in the 1950s. Adding premorbid personality as a psychological variable accounted for some of the variability in character changes after orbital lesions.

Owing to the above qualifications and criticisms, the faith in localizations had been fading since the 1930s. Holists and localizers were still vying for dominance in interpreting the functions of the frontal brain, but strict localizers lost ground (Heygster, 1949). One commentator identified a shift in neuropathology from "cerebrum pathology, psychology, localization doctrine" after the First World War to among other things "comprehensive work of wholeness" [Ganzheitsarbeit] after the Second World War (Birkmayer, 1951 p. V). Correspondingly, researchers accounted for organic character changes without inferring the existence of brain regions for character or morality (De Boer, 1951). Instead of exact localizations, brain scientists more and more aspired to the identification of a "constant relation" (Heygster, 1949 p. 115) between psychological dysfunction and lesion location. This implied the devaluation of the localization doctrine as well as the rise of a new theory of brain functioning. In a review on the psychopathological work on the frontal brain from 1939 to 1955, Häfner (1957) identified a gradual theoretical shift from localizing to

“holistic understandings of the frontal brain syndrome” (p. 238) for the period under review. Häfner proclaimed that Kleist’s localization suggested a false level of accuracy, did not stand the empirical test of the ensuing decades, and was thus to be considered obsolete. Moreover, Häfner stressed that an increment of research in the past years had accumulated a mass of contradictory data, yielding more uncertainty instead of a gain in knowledge. One thing was clear though: the golden times of the classical localization doctrine were over and its simplistic approach to brain functioning had to be denounced. Accordingly, lesions no longer destroyed centers, but disrupted a dynamic neuronal interplay. Assigning an exact location to morality made little sense in this theoretical framework.

In consonance with this shift, researchers steered clear of localizations. Hugo Spatz, who had partly supported Kleist at the meeting in 1936, still pondered the consequences of orbital brain damage because “here man is struck in his intrinsic nature” (Spatz, 1965 p. 237). In a footnote to the respective text, he added that he intentionally avoided the use of the “misguiding term ‘localization’ ” (p. 237). Over the last two decades, Spatz had speculated that the development of the tip of the frontal brain was related to the simultaneous moral improvements in human culture over evolutionary time (Spatz, 1955; Spatz, 1965). In a letter in 1950, Spatz wrote: “I believe that the progressive evolution of the frontal pole of the neocortex will lead to an ethical refinement of humanity” (Peiffer, 2004 p. 627). One can read these statements as Spatz’ speculations on the neuronal basis of morality, though they are explicitly not a localization.

What about Kleist? He had not returned to his localization of morality after 1936. In his last lecture in 1950, he defended neuropathology and its reasoning – that is, the “logic of deletion” (Kleist, 1951). In a solitary paper published after his retirement, he reprinted his controversial localization map featuring the community-I and referred to the orbital brain as “organ of character” (Kleist, 1957 p. 304). Kleist

died in 1960 and simultaneously the era of the last strict localizations of morality was fading out. Testifying to this development are speculations on morality's cerebral underpinnings that appeared at the same time. Wolfgang Klages, a former staff member of Kretschmer in Tübingen and later to become Professor of Psychiatry in Aachen, stated that the "highest mental functions are *tied to* the integrity of the orbital brain" (Klages, 1958 p. 125, my emphasis). The highest mental functions, character, and morality were no longer *localized in* but "*tied to*" the orbital brain. This change of words reflected the redefinition of the localization doctrine: Relations between brain areas and functions had substituted exact localizations. In another publication, Klages put it more succinctly: " 'morality' is neither 'localized' in the frontal brain nor in the interbrain, but rather everything that we consider morality in psychological, sociological, and ethical respect, appears to be mostly dependent upon undisturbed functioning of the frontal and interbrain" (Klages, 1962 p. 53). Morality was no longer localizable.

Conclusion

In this paper, I have documented that based on neuropathological evidence and reasoning researchers attempted to trace morality in the brain in the period from 1930 to 1960. Moreover, I have argued that these attempts took a particular form that reflected the reformulation of the strict version of the localization doctrine in this period. It is important to note that researchers voiced their speculations on morality in the brain mostly in the context of other debates, for example in arguments on the functions of the frontal lobes or the connection between orbital brain damage and criminality. Just as in earlier times, these speculations were scarce and few publications directly addressed morality in the brain. In the period under review, there was no large-scale brain science of morality. As Verplaetse (2009) has suggested, the science was indeed in "crisis" (p. 27) after 1930; a crisis of a specific kind (see below).

In particular, I have shown that Kleist's localization of morality had an afterlife (de Crinis, Meyer, Kretschmer, Thelen, Faust, Walch, Klages); as well as its criticisms. Kleist and his followers were continuously assailed (Bay, Heygster, Gruhle, Häfner). Long familiar arguments echoed through the attacks, mingled with novel criticisms, and changed neuropathologists' conception of morality's neuronal basis. The way that experts deemed morality to be installed in the brain was crucially dependent on the more general theory of brain functioning they adhered to. Kleist's endorsement of a precise localization of the community-I in the orbital brain indicated his adherence to the strict version of the localization doctrine. Later researchers' hesitation to pinpoint morality attested to their skepticism regarding exact localizations. The underlying change in theory expressed itself in De Crinis rejection of an exact localization of morality, Kretschmer's and Schaltenbrand's statistical understanding of localization, Thelen's remarks about brain compensation, Faust's call for considering the temporal dimension in the assessment of lesions, Walch's emphasis on the premorbid personality, and Klages identification of a mere relation between morality and certain brain areas. Though all of these researchers admired Kleist for his seminal work and partly utilized the "logic of deletion", none of them proposed an exact localization of morality.

This absence of exact localizations of morality mirrored a comprehensive shift regarding the general theory of brain functioning. In the first half of the twentieth century, a holistic understanding of brain functioning surfaced, offering an alternative to the strict modularity of the localization doctrine. In addition, the rigidity of the doctrine made it impossible to account for the heterogeneous data (war-inflicted head wounds, tumors, leucotomies) that abounded in the period under review. Researchers gradually realized that the simplicity of the "logic of deletion" disregarded the complexity of the brain. The logic demanded the observation of disturbed functions in conjunction with destroyed regions and then reasoned to regions for intact

functions. The application of this inference to the multitude of lesion cases yielded inconsistent and thus contestable conclusions. Experts questioned classical neuropathological inference and frequently resorted to other methods (e.g. angiography and electroencephalography) that had increasingly become available (Häfner, 1957). The continuing problems (inconsistent results, unsuitable methods, and unjustified generalizations) made theoretical modifications necessary: Neuropathologists started to emphasize the brain's integration, connectivity, and dynamicity. Brain compensation, time-dependent symptoms, and the influences of the premorbid personality challenged the idea that the brain was a bundle of invariant, isolated units. More and more, the strict version of the localization doctrine became untenable: a revision of the theory occurred. Localization doctrine in general had changed considerably since the nineteenth century. As Finger (2010) put it: "Localization theory today is much more complex and dynamic than the localization of sharply demarcated centers envisioned by Broca, Fritsch and Hitzig, and Ferrier" (p. 125). In German neuropathology from 1930 to 1960, aspects of this transition – from a strict version positing isolated centers to an extenuated version proclaiming a division of labor between brain areas– were clearly visible. Accordingly, neuropathologists' notions of morality in the brain changed from moral centers to diversified morality in a dynamic brain. This view resonates through the modern neuroscience of morality in which the search for moral centers is replaced with the quest to identify neuronal networks activated during moral functioning (Schirrmann, 2013b; L. Young & Dungan, 2012). In the past, this transition took place gradually. Kleist's followers simultaneously adhered to the localization doctrine and reformulated and expanded it. The simultaneity of adherence and advancement indicates the passage from one understanding of brain functioning to another. Old and new ideas co-existed for a while. However, by the end of the 1950s exact localizations of morality were outdated. To account for the more or less consistent

connections between orbital and interbrain lesions and moral lapses, experts in this period proclaimed a relation or ties of the brain with morality. The strict version of the localization doctrine was no longer a legitimate method or theory. German neuropathologists devalued the “logic of deletion” as epistemic tool and abandoned the associated ontological claim – that is, morality occupies a specific neuronal spot. Thus, the fate and form of neuropathological attempts to trace morality in the brain in the period 1930 – 1960 has to be understood in the light of the simultaneous redefinition of localization theory.